

# Northumbria Research Link

Citation: Martín-Luengo, Beatriz, Shtyrov, Yury, Luna, Karlos and Myachykov, Andriy (2018) Different answers to different audiences: effects of social context on the accuracy-informativeness trade-off. *Memory*, 26 (7). pp. 993-1007. ISSN 0965-8211

Published by: Taylor & Francis

URL: <https://doi.org/10.1080/09658211.2017.1420196>  
<<https://doi.org/10.1080/09658211.2017.1420196>>

This version was downloaded from Northumbria Research Link:  
<http://nrl.northumbria.ac.uk/32934/>

Northumbria University has developed Northumbria Research Link (NRL) to enable users to access the University's research output. Copyright © and moral rights for items on NRL are retained by the individual author(s) and/or other copyright owners. Single copies of full items can be reproduced, displayed or performed, and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided the authors, title and full bibliographic details are given, as well as a hyperlink and/or URL to the original metadata page. The content must not be changed in any way. Full items must not be sold commercially in any format or medium without formal permission of the copyright holder. The full policy is available online: <http://nrl.northumbria.ac.uk/policies.html>

This document may differ from the final, published version of the research and has been made available online in accordance with publisher policies. To read and/or cite from the published version of the research, please visit the publisher's website (a subscription may be required.)

[www.northumbria.ac.uk/nrl](http://www.northumbria.ac.uk/nrl)





**Different answers to different audiences: Effects of social context on the accuracy-informativeness trade-off**

Journal:	<i>Memory</i>
Manuscript ID	MEM-OP 17-101.R2
Manuscript Type:	Original Paper
Date Submitted by the Author:	14-Dec-2017
Complete List of Authors:	Martín-Luengo, Beatriz; Center for Cognition and Decision Making, National Research University - Higher School of Economics Shtyrov, Yury; Aarhus Universitet, CFIN; Center for Cognition and Decision Making, National Research University - Higher School of Economics Luna, Karlos; University of Minho, School of Psychology Myachykov, Andriy; Northumbria University; Center for Cognition and Decision Making, National Research University - Higher School of Economics
Keywords:	informativeness-accuracy trade off, answering questions, metamemory, uncertainty, conversational pragmatics

SCHOLARONE™  
Manuscripts

Different answers to different audiences: Effects of social context on the accuracy-  
informativeness trade-off

Beatriz Martín-Luengo<sup>1</sup>, Yury Shtyrov<sup>1,2</sup>, Karlos Luna<sup>3</sup>, Andriy Myachykov<sup>1,4</sup>

<sup>1</sup> Centre for Cognition and Decision Making, National Research University Higher  
School of Economics, Russian Federation

<sup>2</sup> Center of Functionally Integrative Neuroscience (CFIN), Department of Clinical Med-  
icine, Aarhus University, Denmark

<sup>3</sup> Psychology Research Centre, School of Psychology, University of Minho, Portugal

<sup>4</sup> Department of Psychology, Northumbria University, United Kingdom

**Author's contact details:**

Dr Beatriz Martín-Luengo, Centre for Cognition and Decision Making, National Research Uni-  
versity Higher School of Economics, 20 Myasnitskaya Ulitsa, 101000 (School of Psychology)  
Moscow (Russian Federation), E-mail: [bmartinluengo@hse.ru](mailto:bmartinluengo@hse.ru)

Professor Yury Shtyrov, Center of Functionally Integrative Neuroscience (CFIN)  
Institute for Clinical Medicine, Aarhus University / Aarhus University Hospital

Running head: CONVERSATIONAL PRAGMATICS AND METACOGNITION 2

Nørrebrogade 44, bldg 10G, 8000 Aarhus, Denmark, E-mail: [yury.shtyrov@cfin.au.dk](mailto:yury.shtyrov@cfin.au.dk)

<http://cfin.au.dk/nedcomm/>

Dr. Karlos Luna, Research Center on Psychology, University of Minho, Campus de Gualtar,

4710-057, Braga, Portugal, E-mail: [karlos.luna.ortega@gmail.com](mailto:karlos.luna.ortega@gmail.com)

Dr Andriy Myachykov, Northumbria University, Newcastle upon Tyne, NE1 8ST, United King-

dom, E-mail: [andriy.myachykov@northumbria.ac.uk](mailto:andriy.myachykov@northumbria.ac.uk)

For Peer Review Only

**Abstract**

Research on conversational exchanges shows that people attempt to optimize their responses' relevance when they definitely know the correct answer (e.g., "What time is it?"). However, such certainty is often unavailable while speakers may still be under social pressure to provide an answer. We investigated how social context influences the informativeness level when answering questions under uncertainty. In three experiments, participants answered difficult general-knowledge questions placed in different social contexts (formal vs. informal). Participants generated their answers, then they were presented with a given context, and decided on the number of alternative responses they wanted to provide (single, with one alternative vs. plural, with several alternatives) and whether the answer should be reported or withheld (report option). Participants reported more answers in the informal context. In the formal context, single answers were preferred, and they were more frequently reported. We conclude that social context influences the level of informativeness in a conversation, affecting achievable accuracy. Our results also show the joint influence of the confidence and the social context on willingness to share information.

**Keywords:** informativeness-accuracy trade off, conversational pragmatics, answering questions, uncertainty, metamemory

**Different answers to different audiences: Effects of social context on the accuracy-informativeness trade-off**

Communicators tailor their messages to the audiences they approach (Clark & Murphy, 1982; Higgins, 1992), as well as modify their interactions and the language used depending on pre-supposed shared common grounds (Isaacs & Clark, 1987) or the communicative goals (Yoon, Koh, & Brown-Schmidt, 2012). For example, we do not report the same amount of information about our weekend activities when asked by a friend (e.g., “I went skydiving wearing a swim suit because I was going to parachute on to a beach”), by our grandmother (“I went to the beach with friends and we had a good time”) or by our boss (e.g. “I went to the beach”). While the first answer is more informative and specific, the others are progressively less informative and more vague. Deciding which information to share becomes even more challenging when we do not have all the information at our disposal. When asked by a friend about the content of the astrophysics article we read, we may not hesitate to state that we did not remember much about it, while in the context of a university exam in astrophysics, we would try to highlight as much as we possibly could remember about the subject. That is, we vary the quality and, especially, the amount of detail of the information we provide (the so-called *informativeness*, see Goldsmith, 2016) depending on the nature of the target audience (Higgins, 1992).

Although there is substantial research on the subject, one aspect that remains largely neglected is the joint effect of the social context and the role of the self-perceived certainty about our memories (retrospective confidence) on the nature of conversational exchanges. In this study we, therefore, pursued three aims: First, we aimed to study how social context influences the amount of conveyed information when the speaker is unsure about the correct answer, i.e., under uncertainty. Second, we aimed

to study the role of the retrospective confidence in this process. Third, we aimed to examine the applicability of a cued-recall memory test (compared with the usual recognition memory tests) to studies of the informativeness-accuracy trade-off.

**Informativeness, pragmatics, and uncertainty**

Research in pragmatics addresses, among other things, the question of how we use language differently in different situations. Specifically, pragmatics is often contrasted to semantics in that it focuses on what is intended rather than on what is said (Gibbs & Moise, 1997). One aspect often discussed in existing literature is the conveyed informativeness or the scalar level of a clause (Huang & Snedeker, 2009; Papafragou & Musolino, 2003). The focus of many scalarity studies is what the *listener* would understand depending on the quantifier used by the *speaker* (some/all, less/much). A typical result is that people may interpret a sentence like “Some presents are ready” as “Not all presents are ready”, even though the use of *some* is in principle consistent with *all*. This assumption derives from the maxim of Quantity (Grice, 1989): The listener expects that the speaker will use the more informative wording in relation to her knowledge. In this case, if all the presents are ready, it makes no sense to use a less informative word like “some”.

In the present study we approached the subject of informativeness from the speaker’s perspective. Specifically, we were interested in the relative change in the informativeness of the speakers’ answers as a function of the social context in which they feature. There are few studies addressing informativeness in natural conversation (e.g., Huang, 2012) and in simple conversational exchanges (i.e., asking for time: Gibbs, & Bryant, 2008; Van der Hernst, Carles, & Sperber, 2002). In the latter studies, participants were approached on the street with a question “Excuse me, do you have the time?” (Gibbs & Bryant, 2008; Van der Hernst, et al., 2002). In both studies,

participants with analogue and digital watches did not provide the exact time (e.g., “it's four fifty-eight”) but an approximate estimate (e.g., “it's almost five o'clock”). Interestingly, when additional pragmatic information was added (e.g., “my watch has stopped”), the percentage of rounded responses such as “almost five o'clock” (less informative than the actual time) decreased for both analogue and digital watch users. That is, participants varied the level of the informativeness in accordance with what they presumed was preferred by the questioner, following the maxim of Quantity (Grice, 1989). However, these results contradict the maxim of Quality (Grice, 1989) that assumes that speakers would strive to be as truthful as possible in their responses. In the example above, to be as truthful as possible is to answer “4:58”. Even though participants could provide an accurate reading, the percentage of rounded responses was never nil, even with the inclusion of pragmatic cues. These results lend support to the relevance theory that states that people provide *optimal* (i.e., adjusted to the questioner's needs) rather than *accurate*, information relevant to the given situation (Wilson & Sperber, 1981; 2004).

Studies using social context manipulations report similar results. For example, Vandierendonck and Van Damme (1988) asked participants to recall a previously heard story. Results showed that participants varied their responses' informativeness depending on the audience (peers, “Martians”, or a public contest). In a recent study by McCallun, Weber, and Brewer (2016, Exp. 1), participants were more willing to provide more specific answers when they were anonymous, suggesting that people are less concerned about providing answers with higher chances of being incorrect when their identity is protected. However, the composition of the audience (high authority, low authority) did not affect the specificity or the responses. The authors explained this



1  
2  
3 result by the relatively young age of the participants (university students) who may not  
4  
5 have differentiated between the two authority levels.  
6

7         In all of the studies reviewed thus far, participants were always certain about the  
8  
9 required responses as the necessary information was available in advance (e.g., in the  
10  
11 studies asking for the time participants knew the exact time). None of these studies  
12  
13 considered the degree of participants' own certainty or uncertainty about the  
14  
15 information they choose to provide. In real life conversations, however, uncertainty  
16  
17 features prominently as we are often unsure about the information we are providing.  
18  
19

20         The nature of the relationship between the social context and uncertainty has not  
21  
22 been comprehensively studied in the literature (although see FeldmanHall, Raio,  
23  
24 Kubota, Seiler, & Phelps, 2015; Yaniv & Foster, 1997). Here, we set out to fill this gap  
25  
26 by investigating how participants change the degree of their responses' informativeness  
27  
28 under uncertainty in varying social contexts. This research sheds light on the processes  
29  
30 underlying speakers' choice on how much information to share with the audience. As  
31  
32 such, it will help us to further our understanding of complex dynamics of real-life  
33  
34 conversational interactions.  
35  
36

37         To address the question of social context effects on informativeness under  
38  
39 uncertainty, we used difficult general knowledge questions. These questions were  
40  
41 placed in two possible scenarios: (1) informal: a conversation with friends and (2)  
42  
43 formal: a job interview. These two contexts differ in terms of their social constraints:  
44  
45 When we are talking with friends, it is more important to be informative than to be  
46  
47 accurate (Yaniv & Foster, 1997) and there is no need to hide a lack of certainty about  
48  
49 the answers. However, in a job interview there is a pressure to present oneself as an  
50  
51 expert in the topic and as accurate as possible, including not reporting incorrect  
52  
53 answers. We are more likely to try to hide our lack of knowledge in formal contexts  
54  
55  
56  
57  
58  
59  
60

(Barrick, Shaffer, & DeGrassi, 2009; Roulin, Bangerter, & Levashina, 2015). Thus, in the informal (“friends”) context we may expect that participants would try to maximize the informativeness since they would be less concerned about the answer’s accuracy, while for the formal context (“job interview”) we expect that participants will strive to provide the most accurate answers. However, how can one measure the informativeness of a given answer? Similarly, how could we combine the measurement of the informativeness with the confidence of remembering?

### **Measuring informativeness and confidence**

The study of informativeness in memory research has been directly linked to response accuracy, and the available results suggest that there is a trade-off between the two: If we want to provide a more accurate response, we tend to decrease its informativeness, and vice-versa (Yaniv & Foster, 1997). In studies of memory and metamemory (the knowledge we hold about memory and our self-awareness of it), there are three procedures to study the accuracy-informativeness trade-off that also provide a measure of the informativeness of an answer: report option, grain-size, and plurality option. In the present study we used a novel variation of the plurality option with the addition of a final phase with the report option. The plurality option approach studies the regulation of the accuracy-informativeness trade-off by manipulating the range of selected answers (Luna, Higham, & Martín-Luengo, 2011). Participants are presented with a question (e.g., “What was the former currency of Spain?”) and several alternatives (e.g., Ducados, Liras, Pesos, Pesetas, Escudos). First, participants are instructed to select one answer (single answer, more informative). Then, participants are requested to add two other alternatives from the original list (plural answer, less informative). Finally, participants have to decide which answer, single or plural, they want to submit as their final response. Participants effectively modify the

1  
2  
3 informativeness of their choice by the strategic selection of the response: single or  
4  
5 plural (Higham, 2013; Luna, et al, 2011; Luna & Martín-Luengo, 2012; Luna, Martín-  
6  
7 Luengo, & Brewer, 2015). Single answers are less likely to be correct but they have a  
8  
9 higher degree of informativeness (i.e., provide more specific information). Plural  
10  
11 answers are more likely to include the correct answer among the alternatives. However,  
12  
13 plural answers are less informative since they include a broader range of alternatives  
14  
15 (Goldsmith, 2016). The *report option* approach (Koriat & Goldsmith, 1994) addresses  
16  
17 the issue of the accuracy-informativeness trade-off by giving participants a binary  
18  
19 option to report their response (informative answer) or to withhold it altogether. A  
20  
21 typical finding is that the accuracy of the final report increases when participants  
22  
23 withhold the answers that have low probability of being correct (Arnold, 2013; Arnold,  
24  
25 Higham, & Martín-Luengo, 2013; Evans & Fisher, 2011; Higham, 2007; Higham, Luna,  
26  
27 & Bloomfield, 2011; Koriat & Goldsmith, 1996; Lueddeke & Higham, 2011).

30  
31 Report option and plurality option procedures are explained by *the dual-*  
32  
33 *criterion satisficing* model (Ackerman & Goldsmith, 2008; Goldsmith, 2016). Both  
34  
35 approaches allow the combined measurement of the informativeness and the confidence  
36  
37 on an answer. In a nutshell, this model states that two criteria need to be satisfied to  
38  
39 provide an answer: confidence and informativeness. Once participants have rated the  
40  
41 subjective likelihood of the alternatives' correctness, the likelihood of the best  
42  
43 alternative is compared against a pre-established criterion (confidence criterion or a set  
44  
45 report threshold). The answer that passes the confidence criterion must also transmit  
46  
47 enough information to be reported. For example, an answer like "Michael is taller than  
48  
49 150 cm" is not appropriate as the range of possibilities is too broad. The  
50  
51 informativeness criterion is supposed to prevent people from providing such general and  
52  
53 uninformative answers. If the answer passes the informativeness criterion, then the  
54  
55  
56  
57  
58  
59  
60

answer is reported; otherwise, the answer is withheld (report option) or coarsened until the criterion is met.

The dual-criterion satisficing model (Ackerman & Goldsmith, 2008) has been used as a theoretical basis in several studies that successfully addressed the accuracy-informativeness trade-off (Higham, 2013; Luna & Martín-Luengo, 2012; Luna et al., 2015; McCallum et. al., 2016). However, the specific constraints and details of the model remain poorly understood (Goldsmith, 2016). The present research attempts to shed light onto two parameters that may affect the dual-criterion satisficing model: (1) The social context (see next section), and (2) the goals (see the introduction to Experiment 2).

### **Cued-recall plurality option and report option**

The plurality option is a flexible approach since it can be used with any type of information. However, it has only been used with multiple choice recognition questions which limits its usefulness to study conversational exchanges because multiple alternatives are not usually provided in a conversation. To circumvent that limitation, the current study employs a novel procedure where participants completed a plurality option procedure and, instead of choosing their answers from a pool of alternatives, they had to generate their own (cued-recall plurality option).

Furthermore, our approach combines the plurality option with the report option. Once participants have chosen the level of informativeness of their answer (single, one alternative; plural, three alternatives), they also have to choose whether they want to report or withhold the selected answer. Combined use of the plurality and the report options will increase the ecological validity of our findings (Neisser, 1985) since in real life we always have the option of not responding. In addition, a study by Koriath & Goldsmith (1994) found increase in accuracy in a free recall test (as opposed to a

1  
2  
3 recognition test) by allowing participants to use the report option. Therefore, regarding  
4  
5 the report option, we expect that participants will sometimes report and sometimes  
6  
7 withhold their answers as a way to regulate the accuracy of their final report, and that  
8  
9 accuracy and confidence will be higher for reported than for withheld answers. This  
10  
11 way, we will obtain information from an experiment with forced answer and, by  
12  
13 introducing the report option, a more close-to-reality information about what happens  
14  
15 when people are free to withhold their candidate answers. Moreover, as stated earlier in  
16  
17 the introduction, both procedures, plurality and report option, were designed to study  
18  
19 the accuracy-informativeness trade-off. However, there is no research that studied the  
20  
21 differences/advantages (if any) of one method over the other (although, see Luna &  
22  
23 Martín-Luengo, 2017). This type of comparison might help future researchers to select  
24  
25 the most appropriate procedure for their purposes.  
26  
27

28  
29 **Cued-recall plurality report option and social context**  
30

31 If we confirm that the procedural innovations introduced here do not negatively  
32  
33 affect the study of informativeness, we will be able to test the effect of social context on  
34  
35 informativeness. Based on the dual-criterion satisficing model (Goldsmith, 2016), and  
36  
37 regarding the social context, we expected a different distribution of responses in each  
38  
39 context. In an informal context, informative answers (single) would be preferred, and  
40  
41 thus more often selected (in the plurality option) and reported (in the report option). In  
42  
43 the formal context, however, we expected that participants would preferentially select  
44  
45 and report the plural answers as a way to avoid incorrect answers and maintain accuracy  
46  
47 high and, therefore, the single answer option would be selected less often. Although  
48  
49 selecting and reporting single answers might give the impression of mastering a topic,  
50  
51 we should take into account that we used difficult questions to create uncertainty in the  
52  
53 answers. Thus, we expected that participants would not be 100% certain about their  
54  
55  
56  
57  
58  
59  
60

answers. As in a job interview, it is reasonable to expect that interviewers will know the correct answers. Providing plural answers guarantees the highest accuracy for difficult questions, but also reduces the cost of providing incorrect single answers and thus of appearing too overconfident in one's knowledge. When the stakes are high, it may better pay to show oneself as prudent and select plural answers than being bold in front of those who may very well know that you are incorrect. In sum, we expected more single answers to be selected, more answers to be reported overall, and more single answers to be reported in the informal than in the formal, context. For these last two hypotheses, we did not expect a parallel pattern of the confidence ratings. The dual-criterion model states that the confidence criterion would differ depending on the situation. Hence, we expect higher confidence ratings for the reported and for the reported single answers in the formal than in the informal context.

## Experiment 1

### Method

**Participants.** Twenty-four volunteers, mostly university students recruited on social media and by posters on campus (23 females, mean age = 20.15,  $SD = 2.91$ ) took part in this experiment in exchange for a small monetary compensation (250• ). The sample size was computed based on the effect size of the main comparison that reflects the success of the regulation of accuracy in the original paper where the plurality option was reported (the comparison of the accuracy between the single selected and rejected answers; Luna, Higham, & Martín-Luengo, 2011, Experiment 1). A power analysis with  $\alpha = .05$  and power = .80 showed that to find a  $d = 1.37$ , 6 participants per cell were enough to find a successful regulation of accuracy.

**Materials and Design.** A normative study was conducted with 25 participants that did not take part in the main experiment and that took part in exchange for a small monetary remuneration (17 females, mean age = 22.00,  $SD = 7.20$ ). Participants were required to answer 86 general knowledge questions. For each question, participants provided three possible answers with a confidence rating for each question ranging from 0% (totally unsure) to 100% (totally sure).<sup>2</sup> The questions were administered in a laboratory using a computer-based questionnaire with self-paced responses. We used general knowledge questions from several sources that were sufficiently difficult so that participants did not always select a single answer, but, at the same time, at a difficulty level that our participants would likely know some of the correct answers. This helped us ensure that some responses would be provided under uncertainty. Forty questions were selected for the main experiment taking into account the criteria that overall accuracy should be lower than .30, and that the provided alternatives had a close relationship with the topic of the question and were reasonable guesses. The actual overall accuracy of the 40 selected questions was:  $M = .23$ ,  $SD = .20$  (see the Supplemental Materials for the full list of the questions selected for the main experiment)<sup>1</sup>.

The experiment had a single independent factor – social context (formal, informal), manipulated within subjects. Dependent variables were: (1) proportion of single and plural answers (for the plurality option), (2) reported and withheld answers (for the report option), (3) confidence, either in the correctness of the answer (for single answers) or in that any of the answers generated was the correct one (for plural answers), (4) accuracy of the different types of answers provided, and (5) reaction times in the selection of the answer (during both the plurality and the report option phases, see below).

**Procedure.** Participants read and signed the informed consent form before taking part in the experiment. Participants were randomly assigned to one of the eight counterbalanced conditions. First, two orders of questions were used, so that half of the participants answered the question in one order and other half in a different one. Second, the context was also counterbalanced per each question (20 questions in each context) such that for half of the participants Question 1 was presented in the informal context, and for the other half it was presented in the formal context. Contexts were randomly assigned to questions, with the only constraint that no more than three successive questions were presented in the same context. This manipulation was preferred over a block design to avoid participants developing a strategy for the given context and then systematically applying it for the rest of the block (but see Experiment 3 where we presented blocks of questions in the same context). Third, we also counterbalanced response key assignment to control for any effect that could be due to handedness: Half of the participants used “p” for single answers and for report answers and “q” for plural and withhold while the opposite was true for the other half. The experiment was programmed using LiveCode (Version 7.1.3) and administered on a desktop computer. After they provided basic demographics, participants read the experimental instructions explaining every phase of the main experiment (for an outline of the procedure, see Figure 1).

After that, both contexts were explained with the help of a background picture. The picture exemplifying the formal (“job interview”) context portrayed three people seated in front of a table with neutral face expressions and dressed formally. The picture of the friends context also portrayed three people, all of them talking and laughing, dressed in casual clothes. Both pictures were accompanied by general descriptive texts



1  
2  
3 stressing that they were either in an important job interview or having a good time with  
4  
5 friends (See Supplemental Materials).  
6

7         Next, participants completed a practice session consisting of three questions  
8  
9 similar to those used in the main experiment. After this, the experimental session  
10  
11 started. For each question, participants first provided a single answer and rated how  
12  
13 confident they were that the answer was correct on a scale from 0% (not confident at  
14  
15 all) to 100% (totally confident) in deciles. Second, they provided two further  
16  
17 alternatives (which, with the single answer, formed the plural answer) and also rated  
18  
19 their confidence that the correct answer was one of the three using the same confidence  
20  
21 scale. The next screen provided a short reminder about the two contexts as well as the  
22  
23 instructions about how to select the final response (single vs. plural) with the keys “p”  
24  
25 or “q”. Then, the context picture for that trial appeared in the middle of the screen. The  
26  
27 single and the plural answers were presented alongside the context picture and  
28  
29 participants had to select one of the two answers as their final response (plurality option  
30  
31 phase). After that, participants saw the context picture again together with the answer  
32  
33 they selected (either single or plural) and they were asked to decide if they would report  
34  
35 or withhold it as their final choice (report option phase), in this case using a computer  
36  
37 mouse. Each individual experimental session lasted approximately half an hour. Finally,  
38  
39 participants were debriefed and dismissed.  
40  
41  
42  
43

44 **Results**

45  
46         One participant was removed because they did not comply with the task. One  
47  
48 question was removed because it turned out to be ambiguous with regard to the  
49  
50 answer’s accuracy (there were two possible correct answers)<sup>3</sup>, thus answers to 39  
51  
52 questions were used in the analysis (producing the total of 897 answers across all 23  
53  
54 participants). Only the questions for which participants provided two or more different  
55  
56  
57  
58  
59  
60

answers (accounting for 758 out of 897 answers) were included in the analyses. Participants reported only one answer when they were certain that it was correct, thus giving no chance to study the informativeness-accuracy trade-off because these answers would be reported regardless of the context. Analyses of the reaction times did not reveal any significant differences between conditions (all  $ps > .05$ ) and will thus not be discussed further. In some analyses the number of participants was lower due to a lack of answers in one of the conditions. We report two-tailed pairwise comparisons obtained with the Student's  $t$  test, and Cohen's  $d$  (henceforth  $d$ ) as a measure of effect size, along with 95% confidence intervals. ANOVAs were not conducted to avoid collinearity problems: If an answer is not selected at single level, it is because it is selected at a plural level. Exploratory analyses showed no effect between the different counterbalance conditions (all  $F_s < 1$ ,  $ps > .05$ ).

**Cued recall test and the plurality option.** The main objective of this experiment was to study conversational informativeness; the main results of the plurality option with a cued-recall test are as such secondary to our objectives and are presented in the Supplemental Materials (section Experiment 1: Cued-recall Plurality option) to the interested reader. Overall, participants did not show a particular preference to choose single answers, something that could have happened because they produced the alternative responses themselves. Also, participants distinguished between single answers with high chances of being correct and low chances of being correct (accuracy of single selected answers was higher than accuracy of single rejected answers), and strategically chose single or plural answers to increase the accuracy of their final response. In sum, our results replicated existing studies using the plurality option, in which alternatives were provided, and showed that the cued-recall test did not affect the sensitivity of the plurality option to study informativeness.

**Cued recall test and report option.** We replicated the results from other studies with the report option and a cued recall tests. There were more reported ( $M = .63$ ,  $SD = .17$ ,  $CI [.56, .70]$ ) than withheld answers ( $M = .37$ ,  $SD = .17$ ,  $CI [.30, .44]$ ),  $t(22) = 3.43$ ,  $p = .001$ ,  $d = 1.44$ . In the same vein, accuracy was higher for reported ( $M = .35$ ,  $SD = .15$ ,  $CI [.29, .41]$ ) than for withheld answers ( $M = .07$ ,  $SD = .09$ ,  $CI [.03, .11]$ ),  $t(22) = 7.78$ ,  $p < .001$ ,  $d = 2.25$ , and confidence followed a similar pattern: higher confidence for reported ( $M = 58.58$ ,  $SD = 16.25$ ,  $CI [51.79, 65.37]$ ) than for withheld answers ( $M = 23.78$ ,  $SD = 17.69$ ,  $CI [16.39, 31.17]$ ),  $t(22) = 9.03$ ,  $p < .001$ ,  $d = 2.04$ . In sum, these results show that the changes introduced in the procedure had no negative effects on the usability of the report option for studying the accuracy-informativeness trade-off.

**Social context and informativeness.** We expected that informativeness would be more relevant in the informal context and that accuracy would be more relevant in the formal context. Therefore, we expected more single answers, more reported answers, and more reported single answers in the informal than in the formal context. Concerning the selection of answers, the results from the plurality option alone showed no reliable differences. In the informal context, participants selected the single answer 46% of the times and in the formal context 55% of the times,  $t(22) = -1.79$ ,  $p = .08$ . However, we confirmed our hypothesis for the report option: there were more reported answers (that is, all the answers reported independently of being selected at a single or plural level) in the informal context ( $M = .76$ ,  $SD = .19$ ,  $CI [.68, .84]$ ) than in the formal context ( $M = .50$ ,  $SD = .21$ ,  $CI [.41, .59]$ ),  $t(22) = 7.45$ ,  $p < .001$ ,  $d = 1.32$ .

The associated confidence ratings showed the opposite pattern: We found a significantly higher confidence in the reported answers in the formal context ( $M = 67.98$ ,  $SD = 16.05$ ,  $CI [61.42, 74.54]$ ) than in the reported answers in the informal

context ( $M = 54.36$ ,  $SD = 17.20$ ,  $CI [47.33, 61.39]$ ),  $t(22) = 3.528$ ,  $p < .001$ ,  $d = .82$ .

This suggests that participants were more willing to report their answers in the informal context and that in the formal context they set a higher confidence criterion. These results illustrate how the social context combined with the confidence in correctness influence the number of reported responses, that is, the information we decide to share with others.

We also expected more reported single answers in the informal than in the formal context (see Table 1; descriptive statistics for all answers are reported for completeness). To test this hypothesis, we computed the relative proportion of single answers that were reported ( $P(r/S)$ ). By computing relative proportions we ruled out any effects of the different proportion of single answers between conditions and incorporated the information from the different proportion of withheld answers. The results showed a higher rate of reported single answers for the informal context ( $M = .80$ ,  $SD = .20$ ,  $CI [.72, .88]$ ) than for the formal one ( $M = .58$ ,  $SD = .27$ ,  $CI [.47, .69]$ ),  $t(22) = 3.37$ ,  $p = .002$ ,  $d = .90$ .

We also confirmed that the confidence in the reported single answers was higher in the formal context ( $M = 42.31$ ,  $SD = 24.01$ ,  $CI [42.31, 24.01]$ ) than in the informal context ( $M = 31.80$ ,  $SD = 18.09$ ,  $CI [24.24, 39.36]$ ),  $t(22) = 3.34$ ,  $p = .003$ ,  $d = .49$ . This result confirms that participants set a higher confidence criterion for reporting, in this case, single answers, in the formal context.

## Discussion

There are three main outcomes of Experiment 1. First, participants chose different levels of informativeness and modified their willingness to report depending on the social context. Second, the context and the confidence in the answer's

correctness affected the amount and the quality of the shared information. Third, the accuracy-informativeness trade-off can be studied with a cued-recall memory test.

Considering that the questions were relatively difficult, we initially expected a larger number of plural answers in the formal context (job interview) as a way of increasing the accuracy of the reports. However, we found that participants withheld more answers in the formal context than in the informal one, suggesting that they preferred to regulate their accuracy by withholding answers, instead of reporting plural answers. In addition, in the informal context participants based their willingness to report on the confidence ratings, independently of the informativeness. On the contrary, in the formal context the differences in confidence were not reflected in a similar pattern at any informativeness' level. In the formal context, participants set a higher criterion to report and reported fewer answers. The prediction that participants would set a different confidence criterion depending on the situation was derived from the dual-criterion satisficing model, but it had never been empirically tested. This is the first evidence showing that different contexts promote different criteria. At the same time, these results also show how the confidence is key to determining the informativeness' level in conversational exchanges.

Another interesting result is that participants did not limit themselves to the first option they produced as they were almost equally likely to choose single and plural answers during the plurality option phase, even though we used a cued recall test. In the informal context (see Table 1) there was indeed no significant differences between the proportion of single and plural reported answers,  $t(22) = -0.41, p = .69$ , but in the formal context there were more single than plural reported answers,  $t(22) = 2.69, p = .01$ , probably due to the implicit requirements of that context.

Regarding the modification of the plurality option procedure to include a cued-recall memory test, our results showed that it is possible to use it to study the accuracy-informativeness trade-off, and also that its efficacy is not affected by the context. This is relevant methodologically since it opens opportunities for the study of the trade-off in other fields and with different materials. It is also theoretically important since it extends the dual-criterion satisficing model.

## Experiment 2

In addition to the social context, the dual-criterion satisficing model (Ackerman & Goldsmith, 2008) also suggests that interlocutors' goals might influence how the confidence criterion is established. Therefore, we introduced goals in Experiment 2 and studied how they might modify our interpretation of the implications provided by a social context and thus alter the pattern of answers. Our theoretical motivation can be illustrated by a real-life situation. Let us consider a conversation with friends that includes a person we want to impress. It is likely that our goal would be to present ourselves in a positive way, trying to be as accurate as possible; in our experiment it would mean predominantly selecting plural answers. Translating this situation into our design, participants in the informal "friends" context and with a substantial informativeness payoff would mimic the pattern of results in the job interview, with fewer answers selected and reported at a single level than with a small payoff.

In Experiment 2, we applied an informativeness payoff approach similar to that used in Goldsmith, Koriati, & Weinberg-Eliezer (2002). In the high-informativeness payoff, the final selection of a correct single answer was rewarded with 5 points and the selection of a correct plural answer with 1 point. In the low-informativeness payoff, participants would earn 1 point for choosing the correct answer. Choosing an incorrect

answer would lead to a loss of 1 point in both single and plural cases. In this low-informativeness payoff, participants typically select the answers at a plural level to maximize accuracy. Goldsmith et al. found that participants in a high-incentive informativeness payoff context were more willing to sacrifice accuracy and reported a significantly higher number of fine answers (equivalent to single answers).

In this experiment, we expected to replicate the overall effect of context from Experiment 1. Regarding the informativeness payoff, we expected more single answers for the high-informativeness than for the low-informativeness condition in both contexts. We also expected more reported answers for the high- than the low-informativeness payoff. In this last case, the associated confidence will not necessarily parallel the direction of the proportion of answers. If there is an effect of the payoff, then the increase in the reported answers might be based on the confidence or not. If the payoff reflects the subjective confidence, then we should find a similar pattern of confidence and proportions, i.e., higher confidence for reported answers in the high- than low-informativeness payoff. However, if the payoff does not affect the distribution of the answers we expect no differences in confidence ratings of both conditions due to the payoff.

Finally, for both contexts we expected more reported single answers in the high-informativeness than in the low-informativeness payoff. Regarding the associated confidence for these answers, we expected either of the two following predictions. If confidence has more weight in the decision than the payoff, we expected higher confidence in the single reported than in the single withheld answers. However, if the payoff has more weight than confidence, then we expected no differences between confidence for single report and single withhold cases, because the payoff may prompt to report some answers that have low confidence.

## Method

**Participants.** Twenty-four volunteers, mostly university students (18 female, mean age = 21,  $SD = 3.89$ ) took part in this experiment in exchange for a small monetary compensation (250• ). Volunteers were recruited on social media and by posters on campus. Participants were also told that their monetary remuneration would be doubled if they could achieve higher score than other participants. The extra-payment was included to boost participants' willingness to pay attention to the payoff.

**Materials and Design.** The same validated questions and the general procedures as in Experiment 1 were used. The main difference was the inclusion of the instructions about the informativeness payoff (see Supplemental Materials). The instructions appeared as a reminder before the plurality option phase, and at the bottom of the picture in both the plurality phase and the report option phase. The design was 2 social context (formal, informal) x 2 informativeness payoff (high, low), manipulated within subjects. The dependent variables were the same as in Experiment 1 but, as reaction times did not produce significant results in Experiment 1, they were not analysed.

**Procedure.** Participants followed the same steps as in Experiment 1 with the following exception: At both stages (plurality option and report option), participants were asked to consider the given social context and to take the informativeness payoff into account. For each question, the context and the payoff were assigned randomly. We had eight counterbalanced conditions as in Experiment 1, but in this case they were created from two orders of questions, two different contexts, in which the answers could be presented, and the two different payoffs that could be applied.

## Results



As in Experiment 1, only the questions, for which participants produced two or more different answers, were included in the analyses. Answers with only one alternative do not allow studying the informativeness-accuracy trade-off since it usually entails that the participant was certain in their answers veracity. These accounted for 828 out of the total of 936 answers. Exploratory analyses showed no effects between the different counterbalanced conditions (all  $F_s < 1$ ,  $p_s > .05$ ). Below we present the analyses that test our hypotheses regarding the effect of the informativeness payoff. We report Analyses of Variance (ANOVA), two-tailed pairwise comparisons using the Student's  $t$  test with Bonferroni correction for multiple comparisons when appropriate ( $p = .012$ ), and  $d$  as a measure of effect size, along with 95% CI. We performed pair-wise comparisons in the absence of a significant interaction (Wilcox, 1987) because we had specific hypotheses based on the dual-criterion model (Goldsmith et al., 2002) that could not be tested otherwise. In addition, this is the first study in which the social context has been manipulated in relation to the informativeness-accuracy trade-off; therefore, it is theoretically important to fully understand the possible relation of this variable to other variables, e.g., the payoff.

We expected more single answers for the high-informativeness than for the low-informativeness condition in both contexts. A 2 (context: informal, formal)  $\times$  2 (informativeness payoff: high, low) repeated measures ANOVA on the mean proportions of the answers selected at a single level revealed no main effects, or interaction (context:  $F(1, 23) = .009$ ,  $p = .923$ ; payoff:  $F(1, 23) = 3.41$ ,  $p = .08$ ; context\*payoff:  $F(1, 23) = 1.21$ ,  $p = .28$ ). To further test our hypotheses, we conducted planned comparisons between the two payoff conditions for each context. The results showed that in the informal context there was no difference in the rate of selection of single answers between the high- ( $M = .50$ ,  $SD = .26$ ,  $CI [.40, .60]$ ) and low-informativeness

1  
2  
3 payoffs ( $M = .46$ ,  $SD = .28$ ,  $CI [.35, .57]$ ),  $t(23) = .57$ ,  $p = .57$ . In the formal context,  
4  
5 however, participants provided significantly more single answers with the high- ( $M =$   
6  
7  $.53$ ,  $SD = .27$ ,  $CI [.42, .64]$ ) than with the low informativeness payoff ( $M = .39$ ,  $SD =$   
8  
9  $.20$ ,  $CI [.31, .47]$ ),  $t(23) = 2.42$ ,  $p = .012$ ,  $d = .61$ . Similar analyses with confidence  
10  
11 showed no differences between high- and low-informativeness payoffs for any of the  
12  
13 two contexts.  
14

15  
16 These results indicate that the inclusion of the informativeness payoff altered the  
17  
18 pattern of responses observed in Experiment 1. Supported by a moderate effect size, the  
19  
20 payoff affected the proportion of single answers only in the formal context. This result  
21  
22 may explain why we did not find differences in the distribution of the answers between  
23  
24 social contexts as in Experiment 1. Additional support for this is that we did not find  
25  
26 parallel confidence ratings with the high- and -low informativeness payoffs. This lack of  
27  
28 differences might suggest that, in formal contexts, participants base their selection of  
29  
30 answers mainly on payoffs, while confidence is of secondary importance.  
31  
32

33  
34 Considering now only the report option, we expected more reported answers in  
35  
36 the condition of high- vs. low- informativeness payoff, regardless of the context. We  
37  
38 performed a 2 (context: informal, formal) x 2 (informativeness payoff: high, low)  
39  
40 repeated measures ANOVA on the mean proportions of reported answers. This analysis  
41  
42 revealed a main effect of the social context,  $F(1, 23) = 12.89$ ,  $p = .002$ ,  $\eta_p^2 = .36$  with no  
43  
44 further significant main effects and/or interactions. There were more reported answers  
45  
46 in the informal context ( $M = .76$ ,  $SD = .19$ ,  $CI [.68, .84]$ ) than in the formal context ( $M$   
47  
48  $= .63$ ,  $SD = .21$ ,  $CI [.55, .71]$ ). However, similar analyses with confidence ratings failed  
49  
50 to reveal any difference between confidence for the reported answers in the informal  
51  
52 context ( $M = 61.73$ ,  $SD = 19.15$ ,  $CI [54.07, 69.63]$ ) and in the formal context ( $M =$   
53  
54  $63.69$ ,  $SD = 18.87$ ,  $CI [56.14, 71.24]$ ,  $p = .23$ . This result indicated that participants  
55  
56  
57  
58  
59  
60

took the payoff into account, at least in the formal context, instead of deriving their decision to report or withhold from the confidence ratings. However, it seems that the effect of the payoff is not sufficiently strong to affect the results of the report option procedure since we did not find a main effect of this variable and we only observed its effect indirectly in the confidence pattern.

Interestingly, while the confidence ratings in Experiment 1 diverged in the opposite direction from the proportion of the reported answers, the results of Experiment 2 did not reveal a similar pattern. One possible explanation of this may be that participants in Experiment 2 provided a higher proportion of answers in the formal context ( $M = .63$ ) than in Experiment 1 ( $M = .50$ ), suggesting that the confidence criterion in Experiment 2 was lower, possibly because of the introduction of the payoff manipulation. In other words, participants might have tried to reduce the increasing cognitive effort resulting from taking into account all the variables in a more complex conversational context by simplifying the situation, e.g., by ignoring some of the variables or highly reducing their weight in the decisions.

Finally, to test the hypotheses that for both contexts there would be more single reported answers in the high-informativeness condition, we performed an ANOVA that revealed a marginally significant effect of the social context,  $F(1, 23) = 4.21, p = .051, \eta_p^2 = .16$ . The descriptive statistics are presented in Table 3 and 4. There were more single reported answers in the informal context than in the formal context. Planned comparisons showed no significant differences between the single reported answers in the high- than in the low-informativeness payoff in either context.

We also computed the relative proportion of single reported answers, that is, the proportion of single answers that were reported. Concerning the hypotheses of more single reported answers for the high-informativeness payoff for each context, the

analyses did not show any significant difference. As there were no differences in proportions, we did not analyse the associated confidence.

Analyses regarding the cued-recall plurality test are presented in the Supplemental Materials, section “Experiment 2: Cued-recall Plurality option for each context and the informativeness payoff”. In short, we found a similar pattern as in Experiment 1, showing that the addition of the payoff did not affect the sensitivity of the procedure to study informativeness.

## Discussion

In summary, Experiment 2 replicated the results of the Experiment 1 with respect to the social context and the use of the plurality option with a cued-recall test. In addition, we found that the payoff only affected the informativeness of the answer at a single/plural level and only in the formal context. Our tentative explanation for this novel finding is that in informal conversational situations social norms are more relaxed and flexible and, as a result, the payoff may be ignored. However, in a more formal conversational situation, we have to follow certain social rules. Thus, the presence of another rule, even if it is not properly a social one such as the payoff, may be perceived as more “natural”; as a result, participants choose to take it into account.

## Experiment 3

Two of the parameters used in Experiments 1 and 2 could be seen as problematic: (1) the information accompanying the images depicting each context and (2) the change of context from question to question. The information provided by the verbose descriptions about how someone could behave in those contexts, while served the purpose of making experimental instructions clear and unambiguous, might have guided participants’ answers. If that were the case, the answer of the participants may

1  
2  
3 have reflected only their following the instructions, and not reacting to each context *per*  
4  
5 *se*. Also, in these two experiments we changed contexts from question to question as a  
6  
7 way to prevent the implementation of a fixed strategy per context. However, this  
8  
9 decision made the manipulation less ecological because in natural conversations we  
10  
11 usually do not change our context from one moment to the next (although it might still  
12  
13 be possible, e.g. when having a conversation with someone in person and at the same  
14  
15 time using instant messaging with another person). To rule out the possible effect that  
16  
17 the instructions and the alternation of contexts might have in Experiments 1 and 2, we  
18  
19 designed another experiment, in which we used the procedure similar to that used in  
20  
21 Experiment 1, but in which the problematic sentences in the descriptions were removed  
22  
23 and the context manipulation was presented in blocks (see below for details).  
24  
25

26  
27 Despite the changes, we expected to replicate the results of Experiment 1.  
28  
29 Regarding the cued recall test and the report option, that is regardless of the context, we  
30  
31 expected participants to report significantly more answers than to withhold them, that  
32  
33 the accuracy of these reported answers would be higher, and that the confidence would  
34  
35 be in line with the accuracy. Regarding the social context and the informativeness, we  
36  
37 expected a similar amount of single answers selected in both contexts but, a  
38  
39 significantly higher number of single selected and reported answers in the informal  
40  
41 context. In this case, confidence should be significantly higher for the single selected  
42  
43 and reported answers in the formal context.  
44  
45  
46  
47

48 **Method**  
49

50 **Participants.** Twenty-eight university students (21 females, mean age = 18.75,  
51  
52 *SD* = 1.74) took part voluntarily in this experiment for a small monetary compensation  
53  
54 (250• ). Participants were recruited on social media and by posters on campus.  
55  
56  
57  
58  
59

**Materials and Design.** Materials and design were the same as in Experiment 1, with the change in the instructions. The descriptions for each context were re-written in a way that would provide no reference to the type of the expected answers just describing the general depicted situation (see Supplemental Materials, section “Experiment 3 Instructions”).

**Procedure.** The procedure was the same as in Experiment 1 except for the following changes. The context was manipulated between four blocks of ten questions each. Participants were randomly assigned to one of the eight counterbalanced conditions: two orders of questions by four possible orders of the 10-question blocks with the social context manipulation (1: informal, formal, informal, formal; 2: formal, informal, formal, informal; 3: formal, informal, informal, formal; 4: informal, formal, formal, informal). Prior to each block, the picture and the description of the context were presented. Also, the picture, but not the verbal description, was presented in one of the corners of the screen in all phases across the block. The answers in the plurality and the report option phases were provided by selecting the appropriate radio button with a mouse. The experiment lasted about 30 minutes.

## Results

Only questions for which participants provided two or more different answers (accounting for 910 out of 1092 answers) were included in the analyses suitable to study the accuracy-informativeness trade-off as has been explained before. We report two-tailed pair-wise comparisons with the Student's *t* test, Cohen's *d* (*henceforth d*) as a measure of effect size, along with 95% confidence intervals. As before, exploratory analyses showed no effect of counterbalancing (all *F*s < 1, *p*s > .05).

**Cued recall test, the plurality option, and the report option.** As for the report option, there were more reported ( $M = .62$ ,  $SD = .18$ ,  $CI [.55, .68]$ ) than withheld

answers ( $M = .38$ ,  $SD = .18$ ,  $CI [.31, .45]$ ),  $t(27) = 3.45$ ,  $p < .002$ ,  $d = 1.30$ . Also, accuracy was higher for reported ( $M = .40$ ,  $SD = .13$ ,  $CI [.35, .45]$ ) than for withheld answers ( $M = .13$ ,  $SD = .08$ ,  $CI [.10, .16]$ ),  $t(26) = 6.42$ ,  $p < .001$ ,  $d = 1.75$ . Finally, the confidence also followed the same pattern of results: higher confidence for reported ( $M = 29.46$ ,  $SD = 16.30$ ,  $CI [23.46, 35.46]$ ) than for withheld answers ( $M = 10.30$ ,  $SD = 11.72$ ,  $CI [6.00, 14.60]$ ),  $t(26) = 5.21$ ,  $p < .001$ ,  $d = 1.35$ . Overall, we fully replicated the pattern of results in Experiment 1.

**Social context and informativeness.** As in Experiment 1, we did not find a preference for the single answers in any of the two contexts: participants selected this type of answer similarly often: 43% in the informal and 45% – in the formal context,  $t(26) = -1.74$ ,  $p = .09$ ,  $d = .65$ . We also replicated the result that there were more reported answers in the informal context ( $M = .69$ ,  $SD = .21$ ,  $CI [.61, .77]$ ) than in the formal context ( $M = .54$ ,  $SD = .21$ ,  $CI [.46, .62]$ ),  $t(27) = 3.44$ ,  $p < .001$ ,  $d = 1.75$ . Regarding the associated confidence, we found that numerically the confidence for the reported answers in the formal context ( $M = 61.89$ ,  $SD = 14.81$ ,  $CI [56.39, 67.39]$ ) was higher than the confidence in the reported answers in the informal context ( $M = 58.02$ ,  $SD = 15.41$ ,  $CI [52.32, 63.72]$ ), although this difference only showed a trend in that direction,  $p = .09$ ,  $d = .26$ . Finally, we did not find any differences in the proportion of reported single selected answers as a function of the context,  $p = .39$ . Therefore, we did not compute the associated confidence.

The full report of the results of the plurality option is presented in the Supplemental Materials. In a nutshell, the results replicated those from previous experiments.

**Discussion**

Experiment 3 showed that the overly verbose descriptions used in Experiments 1 and 2, and the alternation of contexts between individual questions or groups of questions (blocks) did not have any major influence on the results. We fully replicated the pattern of results regarding the usability of the plurality and report option for studying the accuracy-informativeness trade-off: we replicated the main result of Experiment 1, namely the effect of the context on the report option and the preference of participants to use the report option to regulate informativeness. However, we only found a trend, in the same direction of the previous experiments, for the confidence of the reported answers in relation to the contexts.

### General Discussion

The results of the research reported here show, for the first time, that social context and, to some extent, a payoff for informativeness modulate the level of the informativeness of the information exchanged under uncertainty. We conducted three experiments in which difficult questions (validated in a separate study) were used to guarantee a degree of uncertainty in the participants' answers. Questions were placed in two different social contexts: informal ("conversation with friends") and formal ("a job interview"). In Experiment 2, we introduced a payoff manipulation, and in Experiment 3 we ruled out some possible confounds of the procedures used in the first two experiments. Our main conclusions are as follows. First, social context affected the informativeness of the participants' responses as well as the decision about reporting or withholding them. Second, the selected informativeness and willingness to report an answer in informal contexts were driven by the associated confidence, whereas in formal contexts it was mainly driven by implicit norms. Third, goals, such as payoffs, seem to affect conversational exchanges in formal settings. And fourth, it is possible to



use a cued-recall memory test to investigate the accuracy-informativeness trade-off. We will discuss these main findings and address their theoretical implications in more detail below.

Our results showed that the level of informativeness of our answers and the willingness to report vary depending on the social context. The general metamemory model of Nelson and Narens (1990; 1994) states that our answers are affected by both monitoring and control processes. The monitoring process refers to the evaluation of the correctness of our answers. The control process refers to the actions that we will perform or not depending on the specific demands of the social context, our expectations, incentives, goals, and the monitoring evaluation. To date, only one published report used a social context manipulation (authority level) in an accuracy-informativeness trade-off study, and it failed to find any effect (McCallum et al., 2016). Unlike this previous unsuccessful attempt, we found an effect of context – probably because the contexts we used were more easily perceived as different by our participants.

In particular, we found that in the informal context the selection of the level of informativeness and the willingness to report are mainly related to the metacognitive ratings. This is shown in the similar patterns of results between the proportions of the selected answers and the confidence ratings. In the formal context, however, participants based the decisions about the informativeness of the answer and the choice to report on implicit social norms. The results show that the set criterion for the formal context was generally higher than for the informal context, as seen in the confidence ratings; however, participants decided to report single answers instead of reporting more plural answers, likely trying to ensure the highest possible informativeness. That is, the selection of the informativeness and the willingness to report (control processes) were

mainly determined by the implicit norms of that specific context, with a smaller role of the monitoring evaluation.

In Experiment 2, we used an informativeness payoff manipulation in order to investigate how a conversational “goal” might modify the effect of a specific social context. The high payoff for informativeness was hypothesized to increase the amount of single answers in both contexts, but we only found differences in the formal context. Additional confidence results also suggested that our manipulation was successful, although its effect was not as strong as we expected. This was probably because in Experiment 2 the benefit of considering the payoff was linked to the performance of other participants, while it was only measured against the individual performance in previous research. The manipulation we used was the same as in Goldsmith et al. (2002, Exp. 3) except for a critical detail: we offered our participants to double the payment for their participation if they achieved the best score. On the contrary, participants in Goldsmith et al. (2002) received money depending on their individual performance, that is, independent on the performance of others. It is possible that, once they realized that all the questions were difficult, our participants stopped taking into consideration the informativeness payoff because they thought that it was not possible to win the prize. In Goldsmith et al. (2002), participants were not competing against others, so the benefit of at least trying was more evident.

We can conclude that the informativeness level of a conversation might be affected by social context where it is expected to follow rules. Additionally, when there are other norms presented in a conversational exchange, then choices may not be based exclusively on the metacognitive experience, but they also incorporate these norms to decide which answer is preferable. Future research should investigate the possible

differences between answers and confidence patterns in relation to the specific details of the operationalization of the payoff.

**Plurality option cued-recall**

The current study tested the possibility of using the plurality option with cued-recall questions. The first main outcome is that the regulatory processes in memory can be studied with memory tests other than recognition. This is a big step towards a better understanding of memory processes in realistic conversational environments. The recognition tests used until now restrict the generalization of previous results to applied contexts (Sauer & Hope, 2016). Here we showed how this limitation might be overcome with the use of the plurality option procedure combined with a cued-recall test.

The second outcome is that the cued-recall plurality option allows studying conversational exchanges in a controlled environment. While it is very important to study the nature of the conversational exchanges in naturalistic settings, the plurality option cued recall procedure can be successfully used in other, more controlled, set-ups (e.g., neuroimaging studies). However, although in Experiment 3 the block design better approximated the experimental set-up to how every-day conversations usually unfold, there is still a big gap between an interaction through Q&A, as in here,, and a real conversation where both parties request, provide, and share information; future studies are needed to develop controlled approaches to such more realistic situations.

**Social context and informativeness**

An interesting result regarding the effect of context on conversational informativeness is the preferential use of the report option to regulate the level of informativeness. Based on the dual-criterion satisficing model (Ackerman & Goldsmith, 2008), participants were expected to report mainly plural answers in the formal context aiming at increasing their accuracy. However, our participants preferred to increase

accuracy of their answers by withholding responses in this context. That is, they used the report option more often than the plurality option. Despite this, and considering that the questions were difficult, participants did not maximize the use of the report option by withholding the majority of the answers.

One possible explanation for this result is that participants considered this strategy as socially less acceptable. This idea is in line with Ackerman and Goldsmith (2008, Exp. 3). Their results showed that participants used the “don't know” option more often when answering difficult questions than when answering easy questions (45% vs. 30%). This difference, although reliable, was smaller than what the authors expected given that they allowed participants to provide a response coarse enough to include the correct answer. However, it does seem to indicate that admitting our ignorance is perceived as more socially acceptable than reporting extremely uninformative answers. Our results extend these previous findings by demonstrating that the social norms applied to that type of situation (answering difficult questions) vary depending on the context. Specifically, we found more withheld answers for the formal than for the informal context, showing that the context, which the question is presented in, does matter. Studying context effects in more detail to understand the processes that underlie reporting or withholding answers may thus be a fruitful directions for future research.

Another important result was that we did not observe preferential selection of the plural answers in the formal context, considering that all the questions were rather difficult. Participants in the formal context selected the plural answer 45% of the times in Experiment 1, and 49% in Experiment 2, none of these differing significantly from the single-answer selections. One possible explanation for this result is that participants in the job interview (formal) context tried to present themselves as experts in the topic.

From this point of view, selecting a plural answer most of the time is not the most optimal strategy. Indeed, selecting the plural option increases the chances of being correct, but it also implicitly denotes lack of knowledge. This explanation is consistent with the idea that we try to present ourselves as experts in job interviews (Barrick et al, 2009; Roulin et al, 2015). Another explanation is that one may prefer to risk providing a single answer because the alternative might be more embarrassing (Ackerman & Goldsmith, 2008). In this case, potential social embarrassment could overcome the confidence in one's memory for the sake of projecting a more favourable personal image, at least temporarily.

These two strategies, maximizing positive personal impact and minimizing possible negative impact, are directly linked to the notion of the theory of mind (Van Overwalle, & Baetens, 2009). The ability of putting ourselves in other peoples' shoes provides us with a rough estimation of what is expected from us and helps us to respond accordingly. At the same time, this is in line with the approach of the relevance theory that states that people provide the information they believe is relevant for the receiver (Wilson & Sperber, 1981; 2004). Thus, the pattern observed here might result from the beliefs about what the receiver considers relevant and about what we think might help us present ourselves in a better way.

Finally, the fact that participants preferred to select single answers in the formal context can be also explained via the notion of expectancy. When we are talking to our friends (informal context), there are no explicit rules about how to select our answers. However, in a job interview, it is expected that interviewees provide clear and concise answers carrying specific information and not several diverging alternatives to the same question. This third explanation suggests that people's behaviour in the job interview context may not be driven by the need to present oneself as a clever person or the fear of

looking ridiculous, but by the expectancy of what is “normal” in that situation. It has been repeatedly shown that our knowledge is structured as a set of schemata that include the objects and people involved in a given situation (Bartlett, 1932; Lampinen, Copeland, & Neuschatz, 2001; Martín-Luengo, Luna, & Migueles, 2014, 2015) as well a concept of what the expected behaviour is (Bower, Black, & Turner, 1979). Thus, once participants activated their schema of a job interview, they may have tried to behave accordingly. Further research should investigate which of these possible explanations motivate people to select the answer they deem optimal for the given social context.

### **The role of confidence**

Our analyses also highlight the role of confidence in conversation. To the best of our knowledge, this is the first attempt to take into account the confidence in our memories in a study of conversational exchanges. Our results show that selecting an answer is often guided by the confidence we have about its correctness. However, our results also support the idea that in certain circumstances, confidence is disregarded and that answer selection can be explained by the influence of other external demands (e.g., social norms). Specifically, the three experiments showed that in the informal context the confidence ratings parallel the informativeness level selection, as well as the willingness to report (although in Experiment 3 it does not reach significance threshold in the confidence for the report option, so it can be only considered as a trend). However, this is not the case in the formal context, in which the results generally suggest a stronger effect of social norms. Future research will continue to study the situations where we use our metacognitive experiences as the basis for our conversational exchanges.

To summarize, the current study demonstrates how we tailor the information we provide in a conversational setting depending on the audience and the specific social context (e.g., see the optimal relevance theory). Here, we showed how social context influences the informativeness of reported answers by tapping into the metamemory component. As such, the current results document the confidence in our memory as a key factor for our understanding of implicit principles of everyday conversations.

For Peer Review Only

### Acknowledgments

The study has been funded by the Russian Academic Excellence Project '5-100', and supported by the Lundbeck Foundation (Denmark). The authors want to thank Evgenia Lopukhova, Katya Stupina, Elena Kulkova, and Anastasia Liashenko for their assistance in the translation into Russian of the materials and in the recruitment of participants.

For Peer Review Only



**Footnotes**

<sup>1</sup> The original versions of all experimental materials (questions and the description of the contexts in Russian) are available on request.

<sup>2</sup> A full report of the results of the normative study are available on request from the first author.

<sup>3</sup> For the question: “What is the name for the domestic cat with three- or four-coloured pattern?”, in Russian both “calico” and “three-coloured” are correct.

## References

- Ackerman, R., & Goldsmith, M. (2008). Control Over Grain Size in Memory Reporting  
— With and Without Satisficing Knowledge. *Journal of Experimental  
Psychology: Learning, Memory, and Cognition*, 34, 1224-1245. doi:  
10.1037/a0012938
- Arnold, M. M. (2013). Monitoring and meta-metacognition in the own-race bias. *Acta  
Psychologica*, 144, 380-389. doi: 10.1016/j.actpsy.2013.07.007
- Arnold, M. M., Higham, P., & Martín-Luengo, B. (2013). A little bias goes a long way:  
The effects of feedback on the strategic regulation of accuracy on formula-  
scored tests. *Journal of Experimental Psychology: Applied*, 19, 383-402. doi:  
10.1037/a0034833
- Barrick, M. R., Shaffer, J. A., & DeGrassi, S. W. (2009). What you see may not be what  
you get: Relationships among self-presentation tactics and ratings of interview  
and job performance. *Journal of Applied Psychology*, 94, 1394-1411. doi:  
10.1037/a0016532
- Bower, G. H., Black, J. B., & Turner, T. T. (1979). Scripts in memory for text. *Cognitive  
Psychology*, 11, 177-220.
- Clark, H. H., & Murphy, G. L. (1982). Audience design in meaning and reference. In J.  
Le Ny & W. Kintsch (Eds.), *Language and comprehension* (pp. 287–299).  
Amsterdam: North Holland Publishing.
- Evans, J. R., & Fisher, R. P. (2011). Eyewitness memory: Balancing the accuracy,  
precision and quantity of information through metacognitive monitoring and  
control. *Applied Cognitive Psychology*, 25, 501–508. doi:10.1002/acp.1722
- Gibbs, R., & Bryant, G. (2008). Striving for optimal relevance in answering questions.  
*Cognition*, 106, 345-369. doi: 10.1016/j.cognition.2007.02.008

Gibbs, R. W., & Moise, J. F. (1997). Pragmatics in understanding what is said. *Cognition*, 62, 51-74.

Goldsmith, M. (2016). Metacognitive quality-control processes in memory retrieval and reporting. In J. Dunlosky and S. K. Tauber (Eds.) *The Oxford Handbook of Metamemory* (pp 357-385), Oxford, Oxford University Press. doi: 10.1093/oxfordhb/9780199336746.013.28

Goldsmith, M., Koriat, A., & Weinberg-Eliezer, A. (2002). The strategic regulation of grain size in memory reporting. *Journal of Experimental Psychology: General*, 131, 73-95. doi: 10.1037//0096-3445.131.1.73

Grice, H. P. (1989). *Studies in the Way of Words*. Harvard University Press.

FeldmanHall, O., Raio, C. M., Kubota, J., Seiler, M. G., & Phelps, E. A. (2015). The effects of social context and acute stress on decision-making under uncertainty. *Psychological Science*, 26, 1918-1926. doi: 10.1177/0956797615605807

Higgins, E. T. (1992). Achieving "shared reality" in the communication game: A social action that creates meaning. *Journal of Language & Social Psychology*, 11, 107-131. doi: 10.1177/0261927X92113001

Higham, P. A. (2007). No Special K! A signal detection framework of the strategic regulation of memory accuracy. *Journal of Experimental Psychology: General*, 136, 1-22. doi:10.1037/0096-3445.136.1.1

Higham, P.A. (2013). Regulating accuracy on university tests with the plurality option. *Learning and Instruction*, 24, 26-36. doi:10.1016/j.learninstruc.2012.08.001

Higham, P. A., Luna, K., & Bloomfield, J. (2011). Trace-strength and source-monitoring accounts of accuracy and metacognitive resolution in the misinformation paradigm. *Applied Cognitive Psychology*, 25, 324-335. doi:10.1002/acp.1694.

- Huang, Ch. (2012). Referential choice and informativeness in mother-child conversation: A focus on the mother. *Language Sciences*, 34, 319-338. doi:10.1016/j.langsci.2011.10.003
- Huang, Y. T., & Snedeker, J. (2009). Online interpretation of scalar quantifiers: Insight into the semantic-pragmatics interface. *Cognitive Psychology*, 58, 376-415. doi: 10.1016/j.cogpsych.2008.09.001
- Isaacs, E. A. & Clark, H. H. (1987). References in conversation between experts and novices. *Journal of Experimental Psychology: General*, 1, 26-37. doi:10.1037/0096-3445.116.1.26
- Koriat, A. & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General*, 123, 297-316.
- Koriat, A. & Goldsmith, M. (1996). Monitoring and control processes in the strategic regulation of memory accuracy. *Psychological Review*, 103, 490-517
- Lampinen, J. M., Copeland, S. M., & Neuschatz, J. S. (2001). Recollections of things schematic: Room schemas revisited. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 1211- 1222. doi:10.1037//0278-7393.27.5.1211
- LiveCode (Version 7.1.3)[Computer software]. Edinburgh, UK: LiveCode. Downloaded from <https://livecode.com>
- Lueddeke, S., & Higham, P.A. (2011). Expertise and gambling: Using type-2 signal detection theory to investigate differences between regular gamblers and non-gamblers. *Quarterly Journal of Experimental Psychology*, 64, 1850-1871. doi:10.1080/17470218.2011.584631.

Luna, K., Higham, P. A., & Martín-Luengo, B. (2011). The regulation of memory accuracy with multiple answers: the plurality option. *Journal of Experimental Psychology: Applied*, 17, 2, 148-158. doi:10.1037/a0023276

Luna, K. & Martín-Luengo, B. (2012). Improving the accuracy of eyewitnesses in the presence of misinformation with the plurality option. *Applied Cognitive Psychology*, 26, 687-693. doi:10.1002/acp.2845

Luna, K. & Martín-Luengo, B. (2017). Using the regulation of accuracy to study performance when the correct answers is not known. *Scandinavian Journal of Psychology*, doi: 10.1111/sjop.12369

Luna, K., Martín-Luengo, B., & Brewer, N. (2015). Are regulatory strategies necessary in the regulation of accuracy? The effect of direct-access answers. *Memory & Cognition*, 43, 1180-1192. doi:10.3758/s13421-015-0534-2

Martín-Luengo, B., Luna, K., & Migueles, M. (2014). The influence of thematic congruency, typicality and divided attention on memory for radio advertisements. *Memory*, 22, 759-769. doi:10.1080/09658211.2013.826673

Martín-Luengo, B., Luna, K., & Migueles, M. (2015). Effects of interest, thematic congruence, and typicality on memory for television, radio, and press advertisements of new products. *Applied Cognitive Psychology*, 29, 560-572. doi:10.1002/acp.3139

McCallum, N. A., Brewer, N., & Weber, N. (2016). Memorial monitoring and control: How confidence and social and financial consequences affect eyewitnesses' reporting of fine-grain information. *Applied Cognitive Psychology*, 30, 375–386. doi: 10.1002/acp.3212.

- Neisser, U. (1985). The role of theory in the ecological study of memory: Comment on Bruce. *Journal of Experimental Psychology: General*, 114, 272-276.  
doi:10.1037/0096-3445.114.2.272
- Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. H. Bower (Ed.), *The psychology of learning and motivation* (pp. 1–45). New York: Academic Press.
- Nelson, T. O., & Narens, L. (1994). Why investigate metacognition? In J. Metcalfe & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing* (pp. 1–25). Cambridge, MA: MIT Press.
- Papafragou, A. & Musolino, J. (2003). Scalar implicatures: experiments at the semantics-pragmatics interface. *Cognition*, 86, 253-282.
- Roulin, N, Bangerter, A. & Levashina, J. (2015). Honest and deceptive impression management in the employment interview: Can it be detected and how does it impact evaluations? *Personnel Psychology*, 68, 395-444. doi: 10.1111/peps.12079
- Sauer, J., & Hope, L. (2016). The effects of divided attention at study and reporting procedure on regulation and monitoring for episodic recall. *Acta Psychologica*, 169, 143-156. doi: 10.1016/j.actpsy.2016.05.015
- Van Overwalle, F., & Baetens, K., (2009). Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *NeuroImage* 48, 564-584.
- Vandierendonck, A., & Van Damme, R. (1988). Schema anticipation in recall: Memory process or report strategy? *Psychological Research*, 50, 116-122. doi: 10.1007/BF00309211
- Wilcox, R. R. (1987). New designs in analysis of variance. *Annual Review of Psychology*, 38, 29-60. doi: 10.1146/annurev.ps.38.020187.000333

Wilson, D., & Sperber, D. (2004). Relevance Theory. In Horn, L.R. & Ward, G. (eds.)  
*The Handbook of Pragmatics*. Oxford: Blackwell, 607-632.

Wilson, D., & Sperber, D. (1981). On Grice's theory of conversation. In P. Werth (ed.)  
*Conversation and Discourse*. London: Croom Helm. 155-78. Reprinted in  
Kasher (ed.) 1998, vol. 1V, 347-68.

Yaniv, I., & Foster, D. P. (1997). Precision and accuracy of judgmental estimation.  
*Journal of Behavioral Decision Making*, 10, 21-32.

Yoon, S. O., Koh, S., & Brown-Schmidt, S. (2012). Influence of perspective and goals  
on reference production in conversation. *Psychonomic Bulletin & Review*, 4,  
699-707. doi:10.3758/s13423-012-0262-6

Table 1

*Mean (Standard Deviation) [95% Confidence Interval] of the Proportions of Answers in the Plurality Option and Report Option Depending on the Social Context of Experiment 1.*

	Social context			
	Informal		Formal	
	Reported	Withheld	Reported	Withheld
Single	.36 (.17) [.29, .43]	.10 (.09) [.06, .14]	.32 (.22) [.23, .41]	.23 (.15) [.17, .29]
Plural	.39 (.32) [.26, .52]	.14 (.13) [.09, .19]	.18 (.15) [.12, .24]	.27 (.21) [.18, .36]



Table 2

*Mean (Standard Deviation) [95% Confidence Interval] of the Confidence Ratings in the Answers in the Plurality Option and Report Option Depending on the Social Context of Experiment 1.*

	Social context			
	Informal		Formal	
	Reported	Withheld	Reported	Withheld
Single	31.80 (18.09) [24.41, 39.19]	9.18 (14.70) [3.17, 15.19]	42.31 (21.01) [32.5, 52.12]	12.93 (17.18) [5.91, 19.95]
Plural	20.90 (14.47) [14.99, 26.81]	7.54 (11.33) [2.91, 12.17]	18.70 (18.36) [11.20, 26.20]	11.32 (14.31) [5.47, 17.17]

Table 3

*Mean (Standard Deviation) [95% Confidence Interval] of the Proportions of Answers in the Plurality Option and Report Option Depending on the Social Context and the Informativeness Payoff of Experiment 2.*

	Social context			
	Informal		Formal	
	Reported	Withheld	Reported	Withheld
High-informativeness				
Single	.40 (.25)	.09 (.13)	.38 (.27)	.15 (.20)
	[.30, .50]	[.40, .14]	[.27, .49]	[.07, .23]
Plural	.33 (.23)	.17 (.20)	.26 (.21)	.20 (.21)
	[.24, .32]	[.09, .25]	[.18, .34]	[.12, .28]
Low-informativeness				
Single	.39 (.25)	.07 (.12)	.27 (.19)	.12 (.13)
	[.29, .49]	[.02, .12]	[.19, .35]	[.07, .17]
Plural	.38 (.29)	.16 (.17)	.34 (.19)	.27 (.18)
	[.28, .48]	[.09, .23]	[.26, .42]	[.20, .34]

Table 4

*Mean (Standard Deviation) [95% Confidence Interval] of the Confidence Ratings of the Answers in the Plurality Option and Report Option Depending on the Social Context and the Informativeness Payoff of Experiment 2.*

		Social context			
		Informal		Formal	
		Reported	Withheld	Reported	Withheld
High-informativeness					
Single		72.57 (20.80)	16.74 (27.64)	76.67 (20.86)	31.72 (27.70)
		[64.25, 80.89]	[5.68, 27.80]	[68.32, 85.02]	[20.68, 42.73]
Plural		50.31 (27.26)	15.32 (10.13)	54.89 (27.36)	25.78 (20.65)
		[39.40, 61.22]	[11.27, 19.37]	[43.94, 65.84]	[17.52, 34.04]
Low-informativeness					
Single		67.39 (20.90)	20.52 (29.13)	73.72 (23.46)	25.22 (20.89)
		[59.03, 75.75]	[8.87, 32.17]	[64.33, 83.11]	[16.86, 33.58]
Plural		54.79 (22.23)	14.74 (12.20)	51.45 (23.10)	20.78 (16.20)
		[45.90, 63.18]	[9.86, 19.62]	[42.21, 60.69]	[14.30, 27.26]

Table 5

*Mean (Standard Deviation) [95% Confidence Interval] of the Proportions of Answers in the Plurality Option and Report Option Depending on the Social Context of Experiment*

3.

	Social context			
	Informal		Formal	
	Reported	Withheld	Reported	Withheld
Single	.30 (.18) [.23, .37]	.13 (.14) [.08, .18]	.28 (.17) [.22, .34]	.17 (.14) [.12, .22]
Plural	.39 (.23) [.31, .47]	.18 (.13) [.13, .23]	.26 (.20) [.19, .33]	.29 (.21) [.21, .37]

Table 6

*Mean (Standard Deviation) [95% Confidence Interval] of the Confidence Ratings in the Answers in the Plurality Option and Report Option Depending on the Social Context of Experiment 3.*

	Social context			
	Informal		Formal	
	Reported	Withheld	Reported	Withheld
Single	26.74 (18.44) [19.94, 33.54]	7.73 (12.72) [1.73, 13.73]	34.02 (18.93) [27.02, 41.02]	9.47 (11.52) [5.17, 13.77]
Plural	23.62 (16.29) [17.62, 29.62]	11.69 (9.20) [8.29, 15.09]	21.71 (14.07) [16.51, 26.91]	12.26 (9.62) [8.66, 15.86]

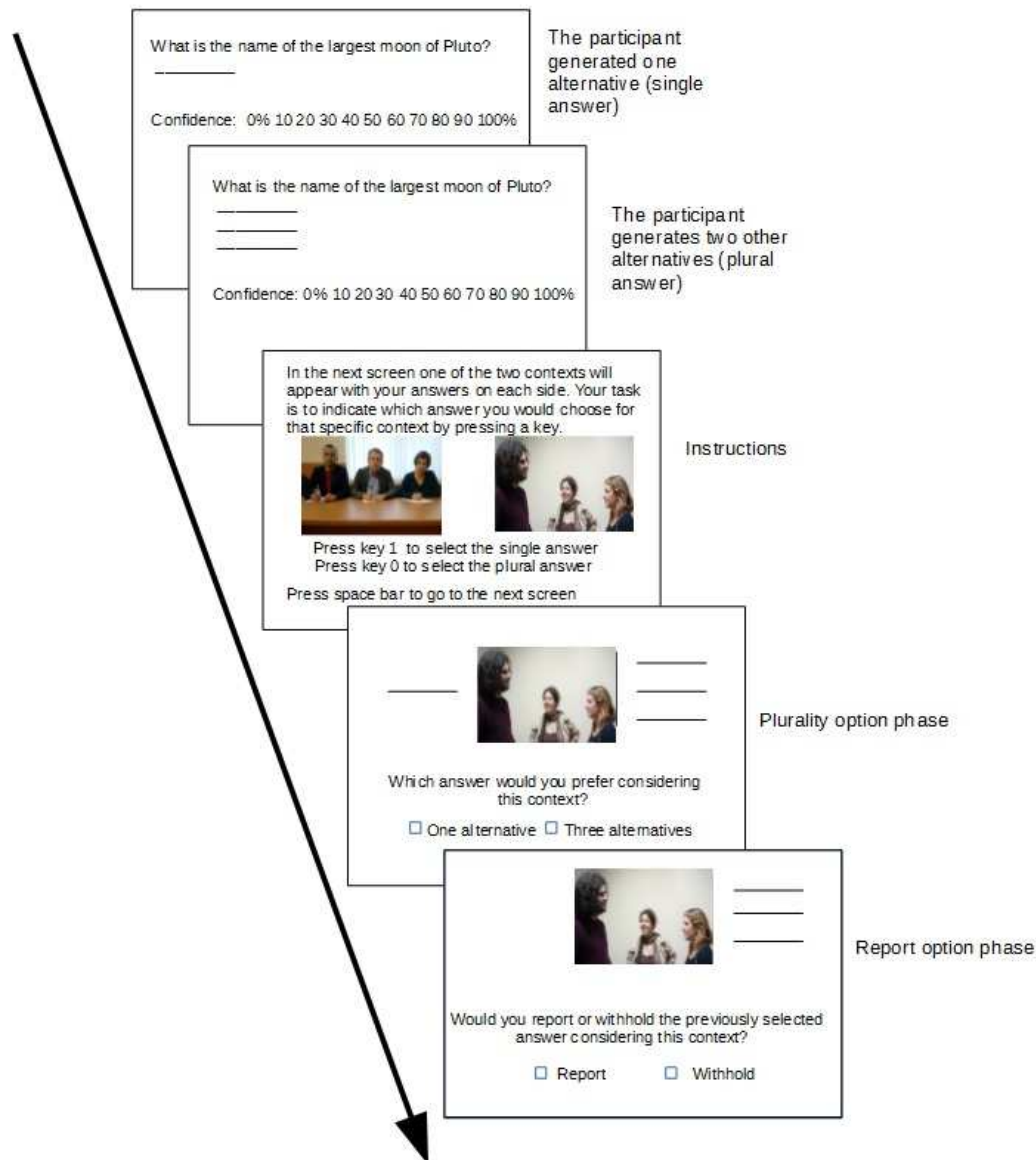


Figure 1. Stages followed for answering the 40 general knowledge questions in Experiment 1. First, participants provided one answer and confidence rating. Second, they added two more alternatives and also provided confidence rating. Then, they completed the plurality and report option procedures.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

**Supplemental Materials**

Different answers to different audiences: Effects of social context on the accuracy-  
informativeness trade-off

## CONVERSATIONAL PRAGMATICS AND METACOGNITION

2

**Translation into English of the original questions in Russian language (correct answer).**

- 1.- What is the term for hitting a volleyball down hard into the opponent's court? (spike)
- 2.- What is the name of the small Japanese stove used for outdoor cooking? (hibachi)
- 3.- What is the name of Pluto's largest moon? (Charon)
- 4.- Which hills separate Scotland and England? (Cheviots)
- 5.- What is the last name of the scientist who discovered Radium? (Curie)
- 6.- In what ancient city were the "Hanging Gardens" located? (Babylon)
- 7.- In which city is Heathrow airport located? (London)
- 8.- What would you be doing if you were spelunking? (exploring a cave)
- 9.- If you have a deficiency of vitamin C, what condition might you get? (scurvy)
- 10.- What is the largest planet in the solar system? (Jupiter)
- 11.- How often enters Halley's Comet orbit the earth? (every 76 years)
- 12.- What was the name of the Zeppelin that exploded in Lakehurst NJ in 1937?  
(Hindenburg)
- 13.- Who was the painter of the Guernica? (Picasso)
- 14.- Who was the painter of the Sistine Chapel? (Michelangelo)
- 15.- What is the name of the first artificial satellite put in orbit? (Sputnik)
- 16.- Who was the discoverer of the vaccination? (Louis Pasteur)
- 17.- What is the capital of Jamaica? (Kingston)
- 18.- What was the name of the ship upon Charles Darwin made his scientific voyage?  
(HSM Beagle)
- 19.- Before the Euro was introduced, which currency was in Portugal? (escudos)



CONVERSATIONAL PRAGMATICS AND METACOGNITION

3

- 20.- What is the name of the astronomical bodies that enter the Earth’s atmosphere?  
(meteors)
- 21.- What is the unit of the sound intensity? (decibel)
- 22.- What is the country with capital in Baghdad? (Iraq)
- 23.- What is the organ that produces insulin? (pancreas)
- 24.- What is the capital of Australia? (Canberra)
- 25.- What is the name for the domestic cats with three- or four-coloured pattern?  
(calico)
- 26.- What is the name of the constellation that looks like a flying horse? (Pegasus)
- 27.- From which country was the composer called Liszt? (Hungary)
- 28.- Where is the Sea of Tranquillity? (Moon)
- 29- What is the main award of the San Sebastian Film Festival? (Golden Shell)
- 30.- What was the earlier name of the city of New York? (New Amsterdam)
- 31- In the ancient India, which was the standard language? (Sanskrit)
- 32- Which number represent 111 in Binary System? (7)
- 33- What is the name of the first man to reach North Pole? (Robert E. Peary)
- 34- What is the name of the Japanese floral art? (Ikebana)
- 35.- What is the colour of the gemstone jet? (black)
- 36.- What is the name of the paradise in the Viking mythology? (Valhalla)
- 37- How long lasted the World War I? (4 years)
- 38- Which is the only South American country that has coast in both Atlantic and Pacific oceans? (Colombia)
- 39- What is the focus of study of herpetology? (amphibians and reptiles)
- 40- Which are the colours of the flag of Finland? (blue and white)

## CONVERSATIONAL PRAGMATICS AND METACOGNITION

4

**Translation into English of the original descriptive texts in Russian language for  
Experiments 1 and 2.**

“The job interview context” (formal)

Imagine that you are in an important job interview. You feel nervous and stressed, but focused on the task. During the interview, you have been questioned on different topics. You noticed that some of your answers could have been more accurate. However, you really need the job, so you keep trying to do your best to show that you are an expert in the topics the interviewers are asking you about. You know that you still have a chance of improving the situation and getting the job.

“The friends context” (informal)

Imagine that you are with some of your close friends and that you are having a good time. You are relaxed, you feel comfortable and pleased to be able to share this good moment with them. You are talking about shared memories and other different topics knowing that you are totally free to openly speak even when you feel that you are not an expert in the specific topic. The atmosphere is relaxed and the important thing is that all of you are together.

CONVERSATIONAL PRAGMATICS AND METACOGNITION 5

**Translation into English of the original descriptive texts in Russian language for**

**Experiment 3.**

“The job interview context” (formal)

"Imagine that you are in an important job interview. You feel nervous and stressed, but focused on the task. During the interview, you have been questioned on different topics. You really need the job, but you know that you still have a chance of getting it."

“The friends context” (informal)

"Imagine that you are with some of your close friends and that you are having a good time. You are relaxed, you feel comfortable and pleased to be able to share this good moment with them. You are talking about shared memories and other different topics, and enjoying that all of you are together."

## CONVERSATIONAL PRAGMATICS AND METACOGNITION

6

**Informativeness payoffs instructions used in Experiment 2.**

“High informativeness payoff instructions”

+5 points if you choose the answer with one alternative and it turns out to be the correct answer

+1 point if you choose the answer with three alternatives and it turns out to be the correct answer

-1 point for choosing an incorrect answer

“Low informativeness payoff instructions”

+1 point for choosing the correct answer

-1 point for choosing an incorrect answer

**Additional analyses.**

**Experiment 1**

**Cued-recall test and plurality option**

The strategic regulation of accuracy is derived from the fact that participants sometimes select the single answer and sometimes the plural answer as their final choice (Goldsmith, 2016). Participants selected single answers 52% of the time and plural answers 48% of the time (difference n.s.), that is, they did not select mainly the plural answer to maximize accuracy, nor did they select the single answer to maximize informativeness.

We also expected, and confirmed, that the accuracy of the selected single answers ( $M = .29$ ,  $SD = .20$ ,  $CI [.21, .37]$ ) was higher than the accuracy of the rejected single answers ( $M = .08$ ,  $SD = .09$ ,  $CI [.04, .12]$ ),  $t(22) = 5.47$ ,  $p < .001$ ,  $d = 1.40$ . This analysis showed that participants were able to distinguish between their correct and incorrect single answers, hence regulating response accuracy.

Our prediction about confidence was also supported: the confidence of the selected single answers ( $M = 50.38$ ,  $SD = 18.36$ ,  $CI [42.88, 57.88]$ ) was higher than the confidence of the rejected single answers ( $M = 31.48$ ,  $SD = 17.40$ ,  $CI [24.37, 38.59]$ ),  $t(22) = 5.96$ ,  $p < .001$ ,  $d = 1.06$ . This result is also in line with other research that states that the discrimination among the correct and incorrect answers is strongly based on metamemory judgments (Koriat & Goldsmith, 1996).

Finally, we also expected that the cued recall test would not interfere with the advantage of the plurality option over a more traditional memory test, in which only one answer is selected. To test this, we compared the accuracy of the final answers, that

## CONVERSATIONAL PRAGMATICS AND METACOGNITION

8

could be either single or plural ( $M = .24$ ,  $SD = .11$ ,  $CI [.20, .28]$ ) against the accuracy of all the single answers ( $M = .19$ ,  $SD = .13$ ,  $CI [.14, .24]$ ). We found the expected increase in accuracy when participants used the plurality option,  $t(22) = 6.54$ ,  $p < .001$ ,  $d = 0.42$ .

**Cued-recall Plurality option for each context**

In the informal context, the accuracy for the selected single answers was higher ( $M = .14$ ,  $SD = .10$ ,  $CI [.10, .18]$ ) than the accuracy for the rejected single answers ( $M = .05$ ,  $SD = .07$ ,  $CI [.02, .08]$ ),  $t(22) = 3.96$ ,  $p < .001$ ,  $d = 1.11$ . In the formal context, we found a similar result. Accuracy was higher in the selected single ( $M = .15$ ,  $SD = .12$ ,  $CI [.10, .20]$ ) than in the rejected single answers ( $M = .02$ ,  $SD = .03$ ,  $CI [.01, .03]$ ),  $t(22) = 5.54$ ,  $p < .001$ ,  $d = 1.50$ . In both cases the discrimination between correct and incorrect answers was based on the associated confidence ratings. In both contexts, we found that the confidence in the selected single answers was higher than the confidence in the rejected single answers (in the informal context, confidence in the selected single answers:  $M = 23.97$ ,  $SD = 10.80$ ,  $CI [19.56, 28.38]$ , and in the rejected single answers:  $M = 13.72$ ,  $SD = 8.87$ ,  $CI [10.10, 17.34]$ ,  $t(22) = 3.87$ ,  $p < .001$ ,  $d = 1.04$ ; in the formal context, confidence in the selected single answer  $M = 26.40$ ,  $SD = 13.33$ ,  $CI [20.95, 31.85]$  and in the rejected single answer  $M = 9.83$ ,  $SD = 6.48$ ,  $CI [4.18, 9.48]$ ,  $t(22) = 6.51$ ,  $p < .001$ ,  $d = 1.58$ ).

Thus, we can conclude that the regulation of the accuracy with a cued-recall test was successful and worked independently of the context.

**Experiment 2: Cued-recall Plurality option for each context and the informativeness payoff**

**Cued-recall plurality option**

The results in this section replicated those from Experiment 1. Participants selected single answers 47% of the time and plural answers 53% of the time (difference ns). This result suggests that participants did not have a strong preference for the plural answers to maximize their accuracy, or for the single answers – to maximize their informativeness. In the formal context, the proportion of the single answer selection was 46% while participants selected plural answers in 54% of the trials. In the informal context participants selected single answers 49% of the time and plural answers 51% of the time.

Next, we assessed whether participants were able to distinguish between their correct and incorrect answers. We found that, pulling all conditions together, the accuracy in the selected single answers was higher ( $M = .38, SD = .22, CI [.29, .47]$ ) than the accuracy in the rejected single answers ( $M = .11, SD = .23, CI [.02, .20]$ ),  $t(23) = 6.15, p < .001, d = 1.50$ . The associated confidence ratings followed the same pattern. Confidence in selected single answers ( $M = 60.50, SD = 19.17, CI [52.83, 68.17]$ ) was higher than confidence in the rejected single answers ( $M = 37.06, SD = 16.70, CI [30.38, 47.74]$ ),  $t(23) = 9.05, p < .001, d = 1.30$ . Thus, participants discriminated their correct and incorrect answers based on their metamemory judgments.

**Social context, payoff, and the cued-recall plurality option**

We also analysed whether participants were able to distinguish between their correct and incorrect answers for each context and payoff. Here, the alpha after

## CONVERSATIONAL PRAGMATICS AND METACOGNITION

10

Bonferroni correction is  $p = .0125$ , because for these analyses there are only four comparisons.

**Informal context.** Participants in the informal context were able to distinguish between their own correct and incorrect answers for both the high and the low payoff. For the high payoff, the accuracy of the selected single answers ( $M = .37$ ,  $SD = .36$ ,  $CI [.21, .53]$ ) was higher than for the rejected single answers ( $M = .04$ ,  $SD = .09$ ,  $CI [0, .08]$ ),  $t(20) = 4.28$ ,  $p < .001$ ,  $d = 1.35$ . This was also true for the accuracy of the selected and rejected single answers for the low payoff (for selected,  $M = .45$ ,  $SD = .35$ ,  $CI [.30, .60]$ ; for rejected,  $M = .08$ ,  $SD = .14$ ,  $CI [.02, .14]$ ),  $t(20) = 5.35$ ,  $p < .001$ ,  $d = 1.34$ . These differences on accuracy were paralleled by their associated confidence ratings. For the high payoff, the confidence in the selected single answers ( $M = 18.49$ ,  $SD = 11.81$ ,  $CI [13.31, 23.67]$ ) was higher than the confidence in the rejected ( $M = 7.22$ ,  $SD = 4.74$ ,  $CI [5.14, 9.3]$ ),  $t(20) = 4.67$ ,  $p < .001$ ,  $d = 1.25$ . The same is true for the confidence in the selected single answers in the low payoff ( $M = 15.55$ ,  $SD = 7.88$ ,  $CI [12.1, 19]$ ) and its rejected equivalent ( $M = 8.21$ ,  $SD = 6.10$ ,  $CI [5.54, 10.88]$ ),  $t(20) = 4.05$ ,  $p < .001$ ,  $d = 1.04$ .

**Formal context.** Results for the formal context were in the same line. Participants in the formal context also distinguished their own correct and incorrect answers for both type of payoffs. For the high payoff, accuracy in the selected single answers ( $M = .34$ ,  $SD = .25$ ,  $CI [.23, .45]$ ) was higher than for the rejected single answers ( $M = .15$ ,  $SD = .25$ ,  $CI [.04, .26]$ ),  $t(21) = 2.77$ ,  $p = .011$ ,  $d = 0.79$ . Likewise, for the low payoff accuracy in the selected answers ( $M = .26$ ,  $SD = .27$ ,  $CI [.14, .38]$ ) was also higher than for the rejected answers ( $M = .14$ ,  $SD = .20$ ,  $CI [.05, .23]$ ),  $t(22) = 2.83$ ,  $p = .010$ ,  $d = 0.59$ . These differences on accuracy were also paralleled with



confidence. For the high payoff, confidence on the selected answers was higher ( $M = 16.76$ ,  $SD = 8.90$ ,  $CI [12.95, 20.57]$ ) than the confidence on the rejected answers ( $M = 6.85$ ,  $SD = 6.15$ ,  $CI [4.22, 9.48]$ ),  $t(21) = 4.20$ ,  $p < .0001$ ,  $d = 1.29$ . For the low informativeness payoff, the confidence on the selected answers was higher ( $M = 13.81$ ,  $SD = 8.17$ ,  $CI [10.32, 17.30]$ ) than the confidence in the rejected answers ( $M = 9.75$ ,  $SD = 8.27$ ,  $CI [6.21, 13.29]$ ),  $t(21) = 2.73$ ,  $p = .012$ ,  $d = 0.49$ . Here, as well as in the previous experiment, we found that the regulation of the accuracy might be studied by a cued-recall test and in different conditions such as different contexts or in the payoffs for informativeness. In both cases, the regulation seems to be grounded in the associated confidence to the selected answers, that is, participants as is stated in the theory, rely on the subjective confidence about the correctness of their answers to decide which answer they would select.

We also compared the accuracy for all the final the answers ( $M = .29$ ,  $SD = .14$ ,  $CI [.23, .35]$ ) with the accuracy for all the single answers ( $M = .23$ ,  $SD = .14$ ,  $CI [.17, .19]$ ). This analysis showed the benefit of using the plurality option instead of a conventional memory test,  $t(23) = 5.87$ ,  $p < .001$ ,  $d = 0.39$ .

In sum, we further validated the results of Experiment 1 by showing that plurality option is a suitable procedure to study the accuracy-informativeness trade off with a cued-recall test even when more than one variable is manipulated.

### Experiment 3

#### Cued-recall test and plurality option

All the results from Experiment 1 were replicated. Participants selected single answers 44% and plural answers 56% of the times, *ns*. Also, accuracy in the selected single answers ( $M = .30$ ,  $SD = .23$ ,  $CI [.21, .38]$ ) was higher than the accuracy of the rejected single answers ( $M = .18$ ,  $SD = .16$ ,  $CI [.12, .24]$ ),  $t(27) = 2.30$ ,  $p = .030$ ,  $d = 0.64$ . The confidence ratings also followed the same pattern of, higher confidence for single selected ( $M = 46.87$ ,  $SD = 15.72$ ,  $CI [41.07, 52.67]$ ) than for single rejected answers ( $M = 34.78$ ,  $SD = 13.74$ ,  $CI [29.68, 39.88]$ ),  $t(27) = 4.75$ ,  $p < .001$ ,  $d = 0.82$ , suggesting that confidence guided the answer selection. Finally, the accuracy for the final answers, regardless of being selected at a single or plural level ( $M = .34$ ,  $SD = .19$ ,  $CI [.27, .41]$ ), was higher than the accuracy of the single answers ( $M = .22$ ,  $SD = .11$ ,  $CI [.18, .26]$ ),  $t(26) = 4.65$ ,  $p < .001$ ,  $d = 0.81$ .

#### Cued-recall Plurality option for each context

In the informal context, the accuracy for single selected ( $M = .27$ ,  $SD = .25$ ,  $CI [.18, .36]$ ) was numerically higher than the accuracy of the single rejected answers ( $M = .17$ ,  $SD = .17$ ,  $CI [.11, .23]$ ) answers, but the difference was only marginally significant,  $t(27) = 1.69$ ,  $p = .10$ ,  $d = 0.44$ . However, in the formal context we found that the accuracy for the single selected answers ( $M = .32$ ,  $SD = .22$ ,  $CI [.24, .40]$ ), was higher than the accuracy for the single rejected answers ( $M = .18$ ,  $SD = .22$ ,  $CI [.09, .26]$ ),  $t(25) = 2.11$ ,  $p = .045$ ,  $d = 0.66$ .

Confidence in either context follows the pattern of Experiment 1. In the informal context, confidence was higher for the single selected answers ( $M = 47.06$ ,  $SD = 18.71$ ,  $CI [40.16, 53.96]$ ) than for the single rejected answers ( $M = 34.87$ ,  $SD = 15.32$ ,  $CI$

CONVERSATIONAL PRAGMATICS AND METACOGNITION 13

[29.17, 40.57]),  $t(27) = 3.65, p < .001, d = 0.72$ . In the formal context, the confidence in the single selected ( $M = 48.61, SD = 14.48, CI [43.21, 54.01]$ ), was higher than for the single rejected answers ( $M = 33.36, SD = 14.89, CI [27.86, 38.86]$ ),  $t(25) = 6.18, p < .001, d = 1.04$ .

For Peer Review Only